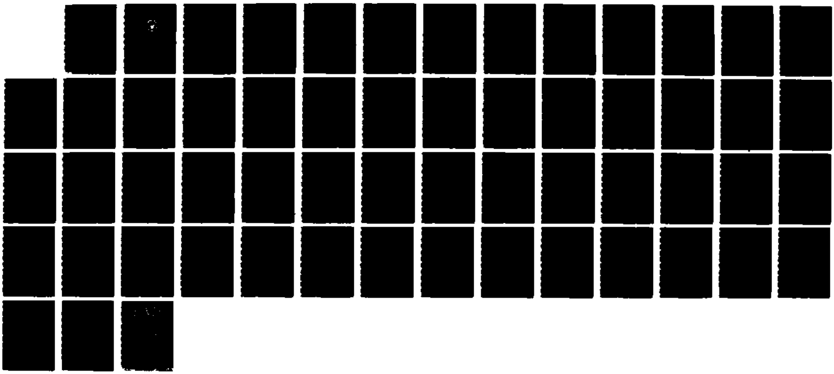
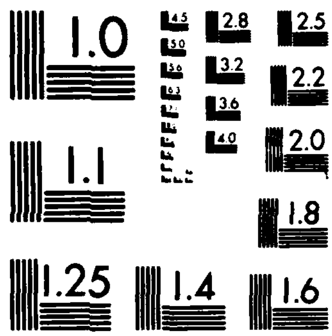
AD-A184 699 TROPICAL CYCLONE INTENSITY PREDICTION BASED ON EMPIRICAL ORTHOGONAL FUNCT (U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA E L WENIGER JUN 87 1/1

UNCLASSIFIED F/G 4/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A184 699

NAVAL POSTGRADUATE SCHOOL

Monterey, California DTIC FILE COPY

THESIS

TROPICAL CYCLONE INTENSITY PREDICTION
BASED ON EMPIRICAL ORTHOGONAL FUNCTION
REPRESENTATION
OF WIND AND SHEAR FIELDS

by

Edward L. Weniger

June 1987

Thesis Advisor Russell L. Elsberry

Approved for public release; distribution is unlimited.

87 9 [illegible]

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A184 679

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b OFFICE SYMBOL (If applicable) Code 63	7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000	7b ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

Tropical Cyclone Intensity Prediction Based on Empirical Orthogonal Function Representation of Wind and Shear Fields.

12 PERSONAL AUTHOR(S)

Weniger, Edward L.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year Month Day)	15 PAGE COUNT
Master's Thesis	FROM ____ TO ____	1987 June	56

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Tropical Meteorology, Tropical Cyclones, Tropical Cyclone Intensity, Empirical Orthogonal Functions

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

An objective technique for predicting 24, 48 and 72 h tropical cyclone intensity is investigated using 1216 cases in the western North Pacific from 1979 to 1983. Potential predictors include conventional storm-related parameters, such as date, intensity, motion and position. Additional potential predictors include empirical orthogonal function (EOF) coefficients of the zonal and meridional components of the environmental wind (250, 400 and 700 mb) and vertical wind-shear (250-400, 400-700, and 250-700 mb) fields. These coefficients represent the synoptic forcing in the vicinity of the storm. The intensity change information is filtered to eliminate data for storms affected by landfall from the sample. The regression equations are verified against a homogeneous sample of Joint Typhoon Warning Center (JTWC) official forecasts, which are also demonstrated to be significantly better (95%

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Russell L. Elsberry	22b TELEPHONE (Include Area Code) 408-646-2373 22c OFFICE SYMBOL 63Es

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted

All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

19. (continued)

confidence) than persistence at all forecast intervals. Regression equations developed using EOF coefficient predictors along with conventional predictors are comparable to the JTWC official forecast, even at 48 and 72 h. The regression equations based on the complete set of predictors have slightly more skill than those based only on conventional predictors. If the regression equations are derived from a smaller sample to allow for an independent test, the results appear to be better in the dependent set, but are degraded in the independent sample. Nevertheless, these independent sample results are comparable in skill to the JTWC forecasts at all intervals. Regression equations generated from three subsets stratified by 12 h old intensity are significantly better than the 48 and 72 h JTWC official forecast.

Approved for public release; distribution is unlimited.

Tropical Cyclone Intensity Prediction Based on Empirical Orthogonal Function Representation of Wind and Shear Fields

by

Edward L. Weniger
Captain, United States Air Force
B.S., Villanova University, 1978

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
June 1987

Author: Edward L. Weniger

Approved by: Russell L. Elsberry, Thesis Advisor

Forrest R. Williams, Second Reader

Robert J. Renard, Chairman,
Department of Meteorology

Gordon E. Schacher,
Dean of Science and Engineering

ABSTRACT

An objective technique for predicting 24, 48 and 72 h tropical cyclone intensity is investigated using 1216 cases in the western North Pacific from 1979 to 1983. Potential predictors include conventional storm-related parameters, such as date, intensity, motion and position. Additional potential predictors include empirical orthogonal function (EOF) coefficients of the zonal and meridional components of the environmental wind (250, 400 and 700 mb) and vertical wind-shear (250-400, 400-700, and 250-700 mb) fields. These coefficients represent the synoptic forcing in the vicinity of the storm. The intensity change information is filtered to eliminate data for storms affected by landfall from the sample. The regression equations are verified against a homogeneous sample of Joint Typhoon Warning Center (JTWC) official forecasts, which are also demonstrated to be significantly better (95% confidence) than persistence at all forecast intervals. Regression equations developed using EOF coefficient predictors along with conventional predictors are comparable to the JTWC official forecast, even at 48 and 72 h. The regression equations based on the complete set of predictors have slightly more skill than those based only on conventional predictors. If the regression equations are derived from a smaller sample to allow for an independent test, the results appear to be better in the dependent set, but are degraded in the independent sample. Nevertheless, these independent sample results are comparable in skill to the JTWC forecasts at all intervals. Regression equations generated from three subsets stratified by 12 h old intensity are significantly better than the 48 and 72 h JTWC official forecast.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I've been blessed abundantly. Many dedicated relatives, friends, educators and associates laid the foundation necessary for me to begin this project. Several people deserve special recognition for contributing directly to it's completion.

Tom and Debbie Schott sponsored me upon assignment to the Naval Postgraduate School. They were supportive friends, especially during those first dismal months in temporary quarters.

LCDR W. E. Wilson derived the EOF coefficients of the zonal and meridional wind fields.

Dr. Ted Tsui of the Naval Environmental Prediction Research Facility provided the intensity data.

The staff, operators and consultants of the W. R. Church Computer Center provided indispensible technical assistance.

Jim Peak helped me acquire data and write computer programs. More important, he instilled confidence . . . computer, and otherwise.

Denis Meanor derived the EOF coefficients of the zonal and meridional wind fields. He helped clear an entangled path by generously sharing his programming and word-processing expertise.

Professor Forrest Williams radiated and fostered a positive attitude; his careful editing was timely, meticulous and constructive.

Dr. Russ Elsberry conceived this project. My mentor, he guided me with his congenial interest, inspired me by his unwavering patience and motivated me through his unbounded energy.

My wife, Dottie, and our children, Eddie and Meaghan, endured much and complained little. Selflessly, they sustained me. I dedicate this work to them with love.

I. INTRODUCTION

A. BACKGROUND

One of the most difficult problems in tropical meteorology is forecasting tropical storm intensity. Numerous models for the prediction of tropical cyclone motion are in operational use at various tropical cyclone centers (Jarvinen and Neumann, 1979; U. S. Command Center/Joint Typhoon Warning Center, 1985). In contrast, there are very few aids for forecasting tropical storm intensity changes in operational use today. Jarvinen and Neumann (1979) suggest this disparity is due to the difficulty in establishing cause and effect relationships for intensity changes. George and Gray (1976) have documented the motion response of the tropical cyclones to environmental "steering" and significant predictor/predictand correlations have been established. Similar well-marked correlations have not been established in the case of intensity changes, at least not for the forecast period beyond 24 h. However, a renewed interest in intensity forecast techniques has recently developed as motion forecasts have improved.

Dvorak (1975) developed an empirical technique based on visual satellite imagery for estimating 24 h intensity changes. The technique was updated (Dvorak, 1982) to incorporate enhanced infrared and digitized satellite imagery, which extended the procedures to nighttime as well as daytime applicability. Unfortunately, this technique is plagued with several limitations and shortcomings. A 24 h forecast is of marginal operational use in support of flight or maritime operations for which more than 24 h leadtime is needed to effectively respond to the threat of a tropical cyclone. In addition, this technique is somewhat subjective; a trained analyst must match current imagery to model storm patterns. Finally, the technique does not handle explosive intensification very well.

Statistical objective intensity forecast techniques based on conventional storm-related data (such as present intensity, latitude, longitude, etc.) were developed by Elsberry *et al.* (1975) for the western North Pacific and Jarvinen and Neumann (1979) for the North Atlantic region. Both studies generated forecast regression equations for periods up to 72 h, rather than the 24 h forecast period characteristic of the Dvorak technique. These techniques basically use a historical sample of storms to develop a

climatology and persistence forecast of intensity similar to the widely used CLIPER track forecast techniques. The basic shortcoming noted in both studies is the characteristic failure of the equations to handle the abnormal case, that is, the rapidly intensifying or decaying storm. Elsberry *et al.* claim that we must improve our ability to recognize the abnormal case if intensity forecasts are to improve. Jarvinen and Neumann suggest we must look beyond the storm-related factors (presumably to environmental influences) to increase our ability to forecast intensity changes. Merrill (1987), who studied tropical cyclone intensity changes in the North Atlantic basin, supports the hypothesis that environmental conditions influence intensity changes of tropical cyclones. However, he concludes the linear relationships are very weak and of little use as objective forecast aids.

The purpose of this study is to demonstrate that empirical orthogonal function (EOF) representations of the zonal and meridional wind fields and of the vertical wind-shear fields can serve as effective predictors of future tropical storm intensity. Shaffer (1982) used an EOF analysis to represent 500 mb geopotential height fields on a grid centered on a tropical cyclone. Shaffer and Elsberry (1982) demonstrated that coefficients from EOF analysis could be used as synoptic forcing predictors in statistical-synoptic track prediction schemes. In a similar study, Wilson (1984) used EOF analysis to represent the 700, 400 and 250 mb wind component fields on a refined grid centered on the cyclone. Wilson (1984) also showed that the coefficients from the wind EOF analysis could be used as synoptic forcing predictors in a statistical track prediction scheme. Schott (1985) used data stratified by past motion to show that the coefficients of the wind EOF analysis could be used as synoptic forcing predictors in a statistical adjustment technique to reduce the systematic errors of a dynamical track prediction model. Meanor (1987) used Wilson's wind component fields to generate EOF fields of vertical wind shear. Using Schott's stratification scheme, Meanor demonstrated that the coefficients from the EOF analysis of wind shear also could be used as synoptic forcing predictors in a statistical adjustment technique to reduce systematic errors of a dynamical track prediction model.

B. OBJECTIVES AND GOALS

The primary objective of this study is to use the existing "conventional" data base; EOF coefficients of wind fields (Wilson, 1984) and wind-shear fields (Meanor, 1987); and selected intensity information to generate useful 24, 48 and 72 h intensity prediction equations for tropical cyclones in the western North Pacific region.

Admittedly, microwave satellite data, cirrus streamer information, sea-surface temperature data, aircraft reconnaissance data and landfall data could also provide meaningful information in intensity forecasting. The goal of this study is to take the first step in developing improved 24, 48 and 72 h intensity prediction schemes for tropical cyclones in the western North Pacific. The eventual goal is to provide an expert system or decision-tree approach, similar to that investigated by Peak and Elsberry (1987) for tropical storm motion, that could be used by the Joint Typhoon Warning Center (JTWC) for operational intensity forecasting.

II. DATA CASE SELECTION

A. DATA DESCRIPTION

The cases in this study are a subset of the cases Wilson (1984) and Meanor (1987) used. These 12-hourly data are for tropical storms in the western North Pacific region for the period from 1979 to 1983. The following restrictions apply to the selection of these cases:

- Tropical storms must be located in the Eastern Hemisphere, east of 100° E with a warning position less than 34.6° N;
- Storm intensity must be at least 18 m/s (35 kt); and
- Zonal and meridional wind components must be available at 700, 400 and 250 mb levels.

A total of 1357 cases meet these requirements.

1. Original Cases (Wilson/Meanor)

a. Conventional Data

The conventional data include observation date/time, storm number and warning positions (current; forecast 24, 48 and 72 h). Additional warning-based information is available as zonal speed, meridional speed and horizontal displacement for three periods: (1) from 12 h prior to observation time until observation time, (2) from 24 h prior to observation time until observation time, and (3) from 24 h prior to observation time until 12 h prior to observation time. Best track positions (current; past 12 and 24 h; and future 24, 48 and 72 h) are also available.

b. Empirical Orthogonal Function Coefficients of Wind and Vertical Wind Shear

The data set for each 12 h case also includes the empirical orthogonal function coefficients of the zonal and meridional wind fields at three levels (Wilson, 1984) and the zonal and meridional shear fields across three layers (Meanor, 1987). The wind information used by Wilson (1984) and Meanor (1987) is from the Global Band Analysis (GBA) operationally generated by the U. S. Navy at the Fleet Numerical Oceanography Center (FNOC). The GBA fields are plotted on a Mercator grid girding the globe from 41° S to 59.8° N, with a grid spacing of 2.5° lat by 2.5° long at 22.5° N and S. The zonal and meridional fields are available from 00 GMT and 12 GMT at the surface, 700, 400, 250 and 200 mb. Surface analyses are based on land observations and ship reports, while upper-air analyses are based on rawinsonde

observations, aircraft reports and satellite-derived cloud motion vectors. Temperature analysis at the intermediate levels are used to couple the winds at adjacent vertical levels via the thermal wind relationship. The 12 h old analysis plus 5% climatology is used as the first-guess field for the current analysis. If no observations are available in a region, the final analysis becomes the previous analysis adjusted towards climatology.

Wilson (1984) defined a relocatable, geographically-oriented grid of 527 points with a fixed zonal and meridional separation of 277.8 km (150 n mi). There are 31 points west to east and 17 points north to south. Thus, the domain is 8334 km (4500 n mi) by 4445 km (2400 n mi). The grid center (row 9, column 16) is coincident with the tropical cyclone center in each case. Wilson used a bi-linear interpolation scheme to extract the zonal and meridional component winds at 700, 400 and 250 mb from the GBA.

Lorenz (1956) first applied empirical orthogonal function analysis to geophysical fields. It has been used regularly to efficiently describe the variability in atmospheric fields. With EOF representations, a large percentage of the variance in a data field can be described by the summation of relatively few orthogonal eigenvectors and their associated coefficients (eigenvalues). This results in a significant reduction in the computer storage space needed to describe synoptic fields, which are ordinarily defined by numerous grid point values.

Wilson generated EOF representations of the zonal and meridional wind fields at three levels (700, 400 and 250 mb) and applied a Monte Carlo approach to select those small sets of rank-ordered eigenvectors and their associated coefficients that describe the signal in the original fields. For this study, the first 35 coefficients of the zonal and the meridional wind fields at each level are available for the 1357 cases. Wilson showed no less than 90% of the variance in all of the zonal wind fields and 82% of the variance in all the meridional wind fields to be explained by the first 35 eigenmodes. The first 25 coefficients of the zonal and meridional wind fields are used as potential predictors in this study.

Following Wilson's methods, Meanor (1987) generated the EOF representations for the zonal and meridional wind-shear fields across three layers: upper (250-400 mb), lower (400-700 mb) and deep (250-700 mb). Meanor also applied a Monte Carlo approach to select those small sets of rank-ordered eigenvectors with their associated coefficients that describe the variance in the wind-shear fields. The first 35 coefficients for the zonal and meridional shear fields across the three layers are

available for the 1357 cases. Meanor showed no less than 80% of the variance in the zonal wind-shear fields and 79% of the variance in the meridional wind-shear fields to be explained by the first 25 and 35 eigenmodes, respectively. The first 25 coefficients of the zonal and meridional wind-shear fields are used as potential predictors in this study.

2. Combined-data cases

In this study intensity data are added to the data set used by Wilson and Meanor. These intensity data are extracted from the Annual Tropical Cyclone Reports for 1979 through 1983 published by the Joint Typhoon Warning Center (JTWC) and include:

- Best track intensity (current, past 12 h, past 24 h, and subsequent 24, 48 and 72 h);
- Warning intensity (current and past 12 h); and
- JTWC official forecast intensity (24, 48 and 72 h).

From these values, best track (past 12 h; and future 24, 48 and 72 h) and JTWC forecast (24, 48 and 72 h) intensity change data are computed. There are 1216 cases in the combined-data cases (Wilson/Meanor data plus intensity data) for use in this study.

B. LAND-OCEAN SORTING

Only storms over the ocean and within the region bounded by the equator, 100° E, 34.6° N and 180° are used in this study. The combined data set is subjected to the following simple land-sea sorting process to separate cases for storms positioned over ocean from cases for storms affected by land.

The bounded region is subdivided into one degree latitude by one degree longitude grid squares, as in Fig. 2.1. If the current or 12 h old position is within a land square or outside the bounded region, the associated 24, 48 and 72 h intensity data are eliminated from the sample. If the current and 12 h old positions are within ocean squares, the positions of the storm at the subsequent forecast times (24, 48 and 72 h) are evaluated. If the position at any of these times is within a land square or outside the bounded region, the intensity change data at that time and all subsequent times are considered unrepresentative and eliminated from the sample.

An example is illustrated in Figure 2.1 based on Typhoon Nelson from March 1982. The current (24/00 GMT), 12 hour old (23/12 GMT) and t+24 h (25/00 GMT) positions of the storm are within ocean grid squares. Because the t+48 h position is located within a land square, the t+48 h and t+72 h intensity change data are

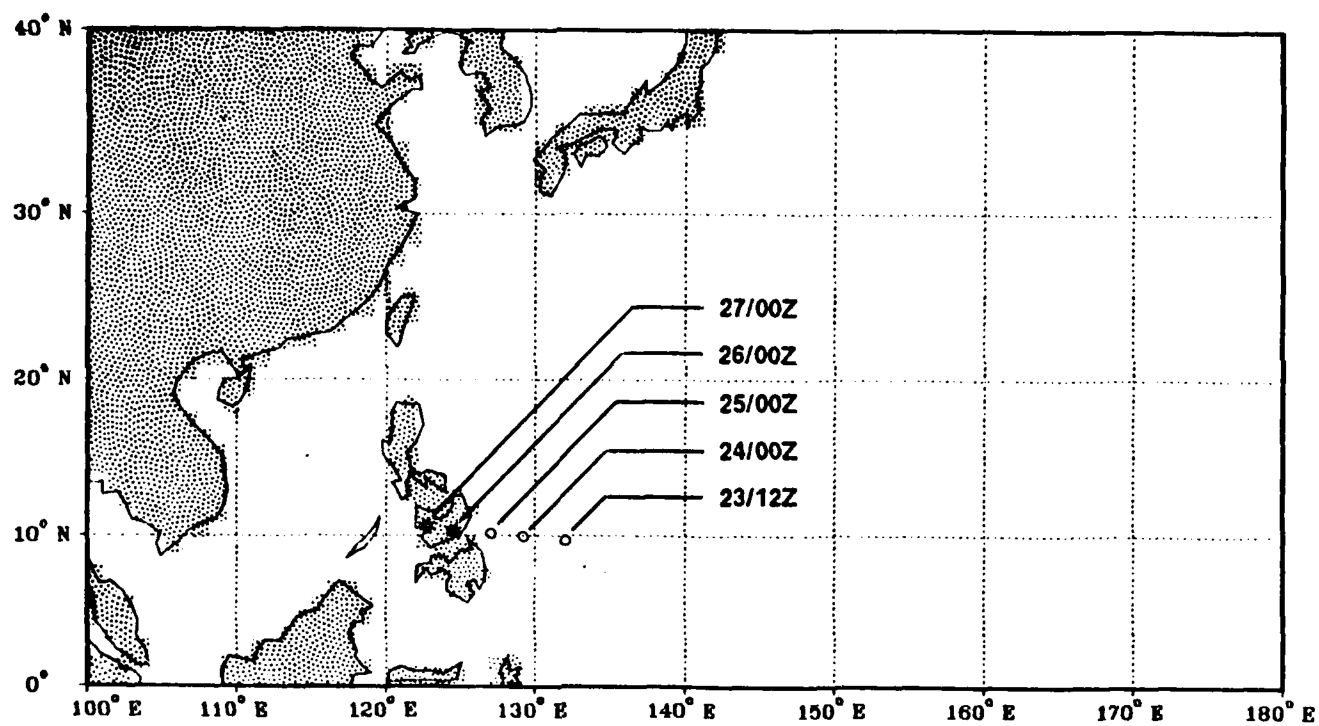


Figure 2.1 Western North Pacific land grid-squares (shaded) and a sequence of Typhoon Nelson best track positions from 12 GMT 23 March 1982 to 00 GMT 27 March 1982.

removed from the sample that is used to derive the regression equations and verify the intensity forecasts.

III. REGRESSION APPROACH

The approach in this study is to use regression analysis techniques to investigate the predictive skill of EOF coefficients of wind and vertical wind shear in forecasting 24, 48 and 72 h changes in tropical storm intensity. The UCLA Biomedical Computer Program (Dixon and Brown, 1985), entitled BMDP2R, is used to select the predictors and to develop the regression model. Tables 1, 2 and 3 are lists of the potential predictors considered for use in the regression equations.

A. POTENTIAL PREDICTORS

The potential conventional predictors are listed in Table 1. The first three predictors (1-3) are the current Julian date and the JTWC warning position (latitude and longitude). The next nine predictors (4-12) describe the storm translation during the past 24 h in terms of the zonal velocity, the meridional velocity and the total displacement. Additional predictors (13-22) in this group include warning intensity data (current, 12 h old and past 12 h change); best track intensity data (current, 12 h old and past 12 h change); and best track position data (current and 12 h old, latitude and longitude).

The second set of potential predictors (23-172), which are listed in Table 2, are the wind-based EOF coefficients generated by Wilson (1984). These represent the external forcing on the cyclone by the environmental winds at three levels (700, 400 and 250 mb). The format used to identify these predictors is CLWNN; where C indicates a wind-based coefficient, L indicates the level (2 for upper, 250 mb; 4 for middle, 400 mb; and 7 for lower, 700 mb), W indicates the zonal or the meridional component wind field (U for meridional, V for zonal wind), and NN is a coefficient number from 1 to 25.

The third set of potential predictors (173-383), which are listed in Table 3, are the wind-shear EOF coefficients generated by Meanor (1987). These represent synoptic forcing upon the storm by vertical differences in the environmental wind through three layers. The format used to identify these potential predictors is SLLWNN; where S indicates a wind-shear coefficient, LL indicates the layer (47 for the lower layer, which is 400 minus 700 mb; 24 for the upper layer, which is 250 minus 400 mb; and 27 for the deep layer, which is 250 minus 700 mb) and NN is a coefficient number from 1 to 25.

B. REGRESSION ANALYSIS

To predict changes in tropical storm intensity over 24, 48 and 72 h, a stepwise regression analysis is used. The BMDP2R program computes estimates of the parameter through a multiple linear regression in a stepwise manner by entering or removing variables one at a time from a list of potential predictors. At each step in the BMDP2R regression analysis routine, the predictor that has the highest partial correlation with the predictand (given the previous selection of predictors) is selected from the remaining set. Consequently, the predictand is the result of a sum of uncorrelated independent variables (Dixon and Brown, 1985).

The F-to-enter value is a function of the number of variables available for selection, their correlation structure and the sample size. In this study, the selection continues until the new predictor does not meet a minimum F-to-enter value of 4.0.

The coefficient of multiple determination (R^2) is a measure of the relationship between the independent and the dependent variables in the regression model and represents the amount of total variance in the predictand that is explained by the independent variables,

$$R^2 = SSR / SSTO = 1 - (SSE / SSTO)\,, \qquad (3.1)$$

where SSR is the regression sum of the squares, SSTO is the total sum of the squares and SSE is the residual sum of the squares. To further restrict the number of predictors in the equations, only those predictors that increase R^2 by at least 0.01 are retained. Finally, an arbitrary limit of ten predictors is set.

TABLE 1

Potential conventional predictors available for the regression analysis.

Number	Name	Description
1	DAYJUL	Julian date
2	LAT	Warning position (latitude)
3	LON	Warning position (longitude)
4	VX0012	Zonal storm speed from -12 h to 00 h (km/h)
5	VY0012	Meridional storm speed from -12 h to 00 h (km/h)
6	V0012	Total storm movement from -12 h to 00 h (km)
7	VX0024	Zonal storm speed from -24 h to 00 h (km/h)
8	VY0024	Meridional storm speed from -24 h to 00 h (km/h)
9	V0024	Total storm movement from -24 h to 00 h (km)
10	VX1224	Zonal storm speed from -24 h to -12 h (km/h)
11	VY1224	Meridional storm speed from -24 h to -12 h (km/h)
12	V1224	Total storm movement from -24 h to -12 h (km)
13	WI00	Warning 00 h intensity
14	WIM12	Warning 12 h old intensity
15	DWIM10	Warning -12 h to 00 h change in intensity
16	BI00	Best track 00 h intensity
17	BIM12	Best track 12 h old intensity
18	DBIM10	Best track -12 h to 00 h change in intensity
19	BLAT	Best track 00 h position (latitude)
20	BLON	Best track 00 h position (longitude)
21	BLTM12	Best track -12 h position (latitude)
22	BLMM12	Best track -12 h position (longitude)

TABLE 2

Potential wind EOF coefficient predictors available for the regression analysis.

Number	Name	Description
23-47	C2U1-25	250 mb wind coefficients derived for zonal modes 1 - 25
48-72	C2V1-25	250 mb wind coefficients derived for meridional modes 1 - 25
73-97	C4U1-25	400 mb wind coefficients derived for zonal modes 1 - 25
98-122	C4V1-25	400 mb wind coefficients derived for meridional modes 1 - 25
123-147	C7U1-25	700 mb wind coefficients derived for zonal modes 1 - 25
148-172	C7V1-25	700 mb wind coefficients derived for meridional modes 1 - 25

TABLE 3

Potential wind-shear EOF coefficient predictors available for the regression analysis.

Number	Name	Description
173-197	S47U1-25	400 minus 700 mb shear coefficients derived for zonal modes 1 - 25
198-222	S47V1-25	400 minus 700 mb shear coefficients derived for meridional modes 1 - 25
223-247	S24U1-25	250 minus 400 mb shear coefficients derived for zonal modes 1 - 25
248-272	S24V1-25	250 minus 400 mb shear coefficients derived for meridional modes 1 - 25
273-297	S27U1-25	250 minus 700 mb shear coefficients derived for zonal modes 1 - 25
298-322	S27V1-25	250 minus 700 mb shear coefficients derived for meridional modes 1 - 25

IV. STUDY METHODS AND VERIFICATION OF RESULTS

A. BASIC METHODOLOGY

The purpose of this study is to investigate the usefulness of empirical orthogonal function coefficients as predictors in an objective forecast scheme of the 24, 48 and 72 h western North Pacific tropical storm intensity. The basic four-part approach is illustrated in Fig. 4.1 and discussed in the following four subsections.

1. Select Data Cases

This study involves the application of regression analysis techniques (Chapter III) to various groupings of the 1216 data cases in the combined-data set (Chapter II). Several groupings of the data are investigated in this study:

- A complete dependent data set (all 1216 cases);
- Dependent-case/Independent-case subsets; and
- Subsets stratified by previous 12 h intensity.

The application of the basic study approach to these data groupings is addressed in Section B of this chapter.

2. Screen Potential Predictors

Because the number of cases in any of the data groupings is small relative to the number of the potential predictors, the potential predictors are screened to determine which are dominant. The predictors are divided into three categories:

- CONV Category - The conventional data listed in Table 1;
- WIND Category - The first 25 EOF coefficients of the zonal and meridional wind fields at three levels (700, 400 and 250 mb) listed in Table 2; and
- SHEAR Category - The first 25 EOF coefficients of the zonal and meridional vertical wind-shear fields across three layers (400-700, 250-400, and 250-700 mb) listed in Table 3.

For each predictand (24, 48 or 72 h intensity change) in each data set/subset, a series of three 10-step regression analyses is performed based on each of the three categories of predictors.

3. Generate Regression Equations

The predictors selected during the screening procedure (a maximum of 30 predictors: up to 10 from each of the three regression analyses) are consolidated. A final 10-step regression is performed using these screened predictors to generate the final equation for the predictand in question.

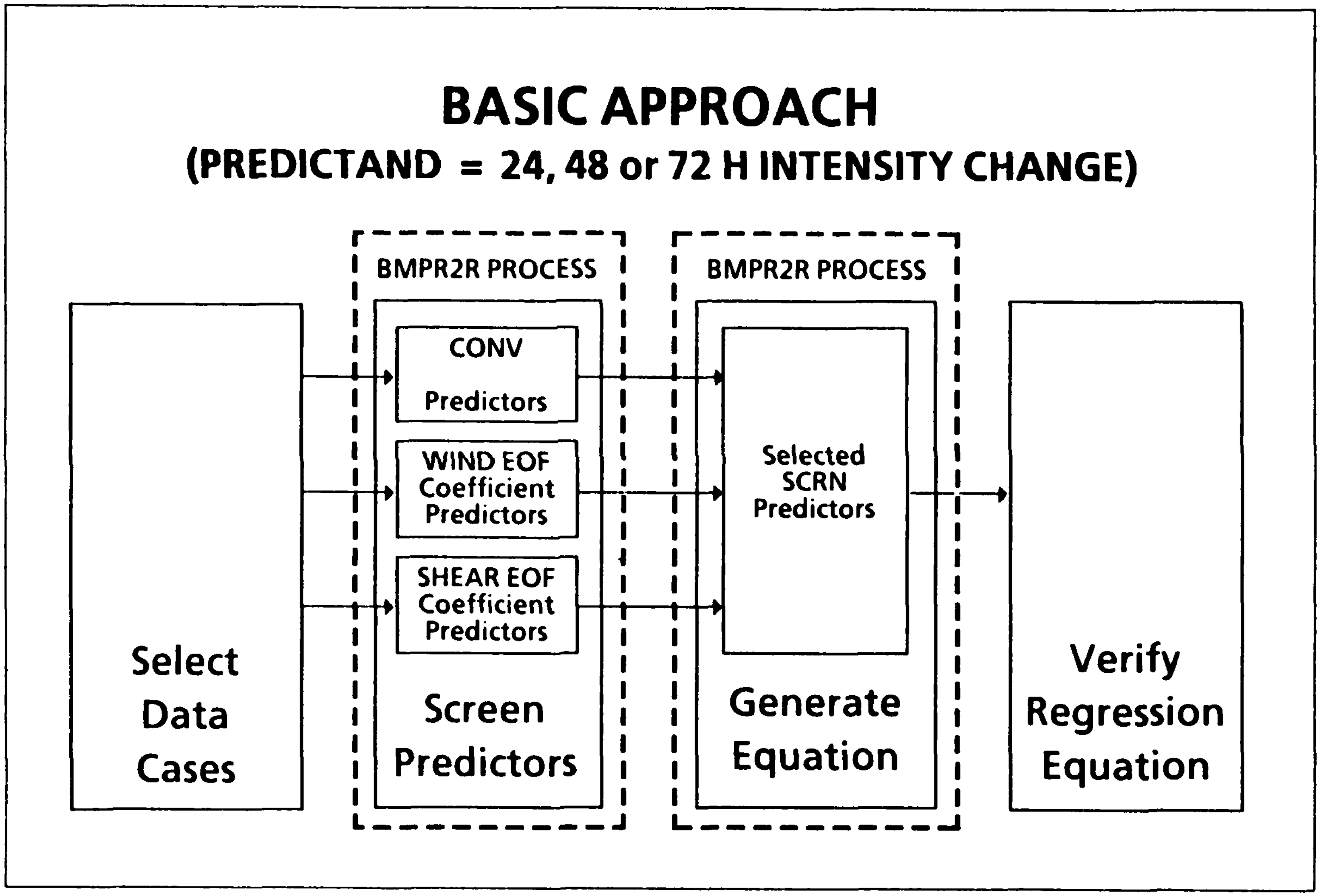


Figure 4.1 Basic study approach (see text).

4. Verify Regression Equations

The regression-derived equations for intensity changes are used to compute forecast intensities at time tt:

$$RItt = WI00 + DItt, \quad (4.1)$$

where WI00 is the warning intensity at observation time, DItt is the regression-derived change in intensity over the forecast interval, and RItt is the regression-derived forecast intensity at verification time tt. The performance of the final regression equations is verified relative to the performance of:

- The JTWC official forecast; and
- A persistence forecast.

The means and standard deviations of the absolute value of the intensity error in the regression, the JTWC and the persistence forecasts are computed and compared. A Student-T test is applied to determine which schemes provide significant improvement at the 95% confidence level.

B. APPLICATION OF THE METHODOLOGY

1. Complete dependent data set

The approach outlined above is first applied to the complete (dependent) data set, i.e., all 1216 cases. Various combinations of best track and or warning predictors are considered for use as the conventional predictors (CONV). The following combination of eight predictors is chosen: date, best track position (current and 12 h old; lat and long), best track intensity (current and 12 h old) and best track past 12 h change in intensity. This combination explains the greatest variance in the intensity change at the three forecast periods and it has the smallest number of missing values.

The predictors selected during the screening process for each of the forecast periods are listed in Tables 4, 5 and 6 for the CONV, WIND and SHEAR category predictors, respectively. The number of potential screened (SCRN) predictors available for each final regression equation is reduced to a maximum of 28 for each predictand: eight conventional predictors, ten wind and ten wind-shear EOF coefficients.

Only three of the eight potential conventional predictors are selected for any one of the three equations (Table 4). The number of predictors selected is limited by

TABLE 4

CONV predictors selected after screening regression on 24, 48 and 72 h best track intensity change (kt) with the complete dependent data set (1216 cases). The numbers indicate the order in which predictors are selected for each equation. The coefficients of multiple determination (R**2) are shown for each equation.

	Forecast Interval		
Predictor	24 h	48 h	72 h
DAYJUL			3
BLAT	3		
BLTM12		3	
BLNM12			2
BIM12	1	1	1
DBIM10	2	2	
R**2	0.33	0.41	0.49

the minimum F-to-enter and change in R^2 requirements applied to the subsequent predictors. The 12 h old intensity, rather than the current intensity, is the first predictor selected at all three forecast periods. At 24 and 48 h, the past 12 h change in intensity is the second predictor selected. This combination of predictors corresponds to a two predictor equation for the extrapolation of the intensity trend. There is no consensus on the additional conventional predictors that are chosen for the three forecast intervals.

Of 150 potential wind EOF coefficient predictors, seven are selected for the 24 h equation and ten are selected for the 48 and the 72 h equations (Table 5). The coefficient of the first eigenmode of the zonal wind at 400 mb (C4U 1) is the first predictor selected in all three equations, while the coefficient of the second eigenmode of the zonal wind at 400 mb (C4U 2) is selected second in the equations at 24 and 72 h (and fourth at 48 h). Wilson (1984) suggests the patterns of modes 1 and 2, which account for the largest variance in the zonal and meridional wind fields, can be interpreted separately as representing particular atmospheric flow patterns. He states

TABLE 5

WIND predictors selected after screening regression on 24, 48 and 72 h best track intensity change (kt) with the complete dependent data set (1216 cases). The numbers indicate the order in which predictors are selected for each equation. The coefficients of multiple determination (R**2) are shown for each equation.

	Forecast Interval		
Predictor	24 h	48 h	72 h
C2U14	7		
C2V25		6	
C4U 1	1	1	1
C4U 2	2	4	2
C4U16		9	10
C4U20		7	
C4V12	6		4
C7U 7	4	3	6
C7U17		8	8
C7U24			9
C7V 3	3	2	7
C7V11	5		
C7V13			5
C7V14		5	3
C7V16		10	
R**2	0.26	0.32	0.36

that the complexity of the eigenvalues makes it difficult to associate higher order modes for any of the fields with any observable atmospheric patterns. If the 400 mb conditions can be assumed to represent the mean flow through the depth of the troposphere, the first coefficient of the zonal wind is indicative of the mean zonal environmental flow. A positive value (related to easterly flow) generally may be associated with storm development, while a negative value (related to westerly flow) would imply recurvature and associated weakening.

TABLE 6

SHEAR predictors selected after screening regression on 24, 48 and 72 h best track intensity change (kt) with the complete dependent data set (1216 cases). The numbers indicate the order in which predictors are selected for each equation. The coefficients of of multiple determination (R**2) are shown for each equation.

Predictor	Forecast Interval 24 h	48 h	72 h
S47U 1		1	1
S47U 2	2	6	
S47U 3			5
S47U 4		4	
S47U 8	3		
S47V 3		3	
S47V15			2
S24U 1	1		7
S24U 2	4		
S24V20	6		
S27V 1		2	6
S27V 4	5		
S27V 9			4
S27V11		5	
S27V17			8
S27V19			9
S27V22			3
R**2	0.16	0.16	0.22

Of the 150 potential wind-shear coefficients, six are selected for the 24 and 48 h equations and nine for the 72 h equation (Table 6). In contrast to the selected conventional and wind EOF coefficient predictors, the wind-shear EOF coefficients are less consistent in time. None of the wind-shear coefficients are selected for all three equations (24, 48 and 72 h). Only four wind-shear predictors (S47U 1, S47U 2, S24U 1 and S27V 1) are selected for two of the three equations.

Notice that the explained variance increases with increasing forecast interval for all three categories of potential predictors. The conventional predictor equations account for the most explained variance, while the wind-shear EOF predictor equations account for the least explained variance. Of the nine equations, the conventional predictor equation for the 72 h forecast intensity explains the most variance.

Before combining the screened predictors and doing a final regression using these selected screened predictors, the performance of the equations derived from the three separate categories of predictors is investigated (Table 7). Analysis of Table 7 shows that the mean intensity forecast error and the standard deviation of the intensity forecast error increase as the forecast interval increases for all schemes (JTWC; CONV, WIND and SHEAR predictor). For all forecast intervals, the equations generated using only the best track conventional predictors perform better (have smaller average absolute errors) and are more consistent (have smaller standard deviations in the average absolute error) than the equations generated using only wind EOF coefficient predictors, which perform better and are more consistent than the equations generated using only wind-shear EOF coefficient predictors. Although the official JTWC intensity forecast errors are smaller than all the regression-derived equations at 24 h, the best track conventional predictor equations perform better and are more consistent than JTWC at 48 and 72 h. Recall that these results are for a dependent sample. Presumably, even more accurate predictions are possible if all three categories of screened predictors are included.

The three screened-predictor regression-derived equations for the 24, 48 and 72 h intensity change and the coefficient of multiple determination (R^2) for each predictor are indicated in Table 8. For all forecast intervals, the regression process terminates before ten predictors are selected, because the F-to-enter values or the amount of variance explained by the subsequent predictors are too small for further stepping. Only 4, 6 and 5 predictors are selected at 24, 48 and 72 h, respectively. As suggested by Table 4, the 12 h old best track intensity is the first predictor chosen for all three forecast intervals (24, 48 and 72 h). This observation prompted a later stratification of the data (to be discussed in Section B.3 below) based upon 12 h old best track intensity. Several wind EOF coefficient predictors appear in the screened predictor equations. In fact, C4U 1 and C4U 2 are among the top four predictors in all three equations. No EOF coefficients of wind shear are chosen. As Meanor (1987) suggests, perhaps this is due to the close relationships between the wind and wind-

TABLE 7

Verification of JTWC and regression-derived (CONV-, WIND- and SHEAR-predictor) forecasts of 24, 48 and 72 h tropical storm intensity (kt) for the complete dependent data set (1216 cases) based on land-filtered and homogeneous samples.

JTWC Forecast Intensity

	Cases	Avg Abs Error	Std Dev
24 h	886	13.1	11.3
48 h	651	21.3	16.6
72 h	462	24.5	19.0

Best Track Conventional Predictors (CONV)

	Cases	Avg Abs Error	Std Dev
24 h	886	13.5	11.5
48 h	651	20.9	15.6
72 h	462	22.6	17.2

Wind EOF Coefficient Predictors (WIND)

	Cases	Avg Abs Error	Std Dev
24 h	886	14.6	12.3
48 h	651	22.5	16.9
72 h	462	25.6	19.5

Shear EOF Coefficient Predictors (SHEAR)

	Cases	Avg Abs Error	Std Dev
24 h	886	15.4	12.8
48 h	651	24.5	18.7
72 h	462	28.2	20.6

shear synoptic forcings. After a wind EOF coefficient predictor is selected, the wind-shear EOF coefficients that are highly correlated with it will not be selected.

TABLE 8

Regression equations for the change in intensity (kt) at 24, 48 and 72 h using the complete dependent data set (1216 cases). Parenthetical values indicate the order in which the screened predictors are selected for each equation. The coefficients of multiple determination (R**2) are shown.

	Forecast Interval		
	24 h	48 h	72 h
Y-Intercept	15.64	-0.97	48.71
Predictor			
BLNM12	-	0.28 (6)	-
BIM12	-2.24 (1)	-0.57 (1)	-0.78 (1)
DBIM10	0.46 (2)	-	-
C2U14	-	-1.95 (5)	-
C4U 1	0.54 (3)	1.03 (2)	1.32 (2)
C4U 2	0.38 (4)	0.71 (3)	1.11 (4)
C4V12	-	-	1.31 (5)
C7U 7	-	-1.36 (4)	-1.71 (3)
Cases	886	684	512
R**2	0.39	0.51	0.56

Interestingly, the regression-derived equations explain a larger percentage of variance in the predictand with increasing forecast interval. This is a favorable result because the objective is to provide forecast guidance at 48 and 72 h. However, notice that the maximum value of explained variance (at 72 h) is only 56%; i.e., 44% is still unexplained.

The performance of the SCRN predictor equations relative to the performance of homogeneous samples of persistence, JTWC and CONV predictor forecasts is illustrated in Table 9. At all forecast hours, the smallest mean absolute errors are associated with the regression-derived intensity forecasts generated using the SCRN predictor equations. In addition, the standard deviations of mean absolute errors associated with these equations are the smallest, which indicates more consistent forecasts.

TABLE 9

Verification of persistence, JTWC and regression-derived (CONV- and SCRN-predictor) forecasts for 24, 48 and 72 h tropical storm intensity (kt) for the complete dependent data set (1216 cases) based on land-filtered and homogeneous samples.

Persistence Forecast

	Cases	Avg Abs Error	Std Dev
24 h	886	17.2	13.7
48 h	651	28.1	19.8
72 h	462	33.7	23.9

JTWC Forecast

	Cases	Avg Abs Error	Std Dev
24 h	886	13.1	11.3
48 h	651	21.3	16.6
72 h	462	24.5	19.0

Regression-Derived Forecast
Best Track Conventional Predictors (CONV)

	Cases	Avg Abs Error	Std Dev
24 h	886	13.5	11.5
48 h	651	20.9	15.6
72 h	462	22.6	17.2

Regression-Derived Forecast
Selected Screened Predictors (SCRN)

	Cases	Avg Abs Error	Std Dev
24 h	886	13.0	11.2
48 h	651	19.5	14.8
72 h	462	21.3	16.1

Notice that the JTWC forecast performs better than persistence, particularly at 72 h. Student-T significance tests indicate that the JTWC forecast is better than

persistence (95% confidence level) at all forecast hours. Because the intensity observations and forecasts are rounded to the nearest 5 kt value at each end of the change interval, discretization errors result. Therefore, no scheme is likely to perform with a minimum error of less than 10 kt. Although the 13 kt mean absolute error of the JTWC official forecast at 24 h is relatively good, this error approximately doubles by 72 h.

The results of the CONV predictor equations (the basis for existing intensity forecast schemes) are repeated from Table 7 for comparison with the SCRN predictor equations. Notice that the additional contribution of synoptic predictors (wind and wind-shear EOF coefficients) in reducing the mean absolute error is small (0.5 kt at 24 h, 1.4 kt at 48 h and 1.3 kt at 72 h). Comparison of Table 7 with Table 9 suggests that much of the variance explained by the WIND (or SHEAR) predictors is already contained in the selected CONV predictors. Nevertheless, the synoptic forcing represented by the EOF coefficients does lead to significant intensity forecast improvements at 48 and 72 h in this dependent data sample.

2. **Dependent-case/Independent-case subsets**

The above results based on the dependent sample may be overly optimistic, because the verification cases were used to derive the regression equations. Thus, the data cases were subdivided into dependent-case and independent-case subsets:

- To investigate the effect reducing the sample size would have on the regression-derived equations for 24, 48 and 72 h intensity change, and
- To investigate the predictive skill of the dependent-case regression-derived equations when applied to an independent-case data subset.

The independent-case subset of 405 cases is constructed by selecting every third case in the complete data sample. The dependent-case subset is the remaining 811 cases.

The basic approach in Fig. 4.1 (Section A above) is applied to the dependent sample. The resulting equations are listed in Table 10 for the 24, 48 and 72 h intensity change. For each forecast interval, the predictors selected first and explaining the largest percentage of the variance in the predictands in Table 10 are common to the equations derived using the complete dependent set (Table 8). In the 24, 48 and 72 h regression equations, the selection sequence is common between the two data sets for the first 3, 4 and 3 predictors, respectively. More predictors are selected for the dependent-case subset equations than the complete dependent set (6, 7 and 6 versus the 4, 6 and 5 at 24, 48 and 72 h). The dependent-case subset equations explain slightly more variance (0.42, 0.52 and 0.59 compared to 0.39, 0.51 and 0.56) than the complete

dependent set equations. This is expected, since the sample sizes are smaller; i.e., they contain less of the natural variability of the ensemble of possible cases. However, adding more predictors may not lead to better predictions in an independent test.

TABLE 10

Regression equations for the change in intensity (kt) at 24, 48 and 72 h using the dependent-case data subset. Parenthetical values indicate the order in which the SCRN predictors were selected. Asterisks indicate common predictors with the corresponding equations for the complete dependent set in Table 8. The coefficients of multiple determination (R**2) are shown.

	Forecast Interval					
	24 h		48 h		72 h	
Y-Intercept	26.27		35.61		46.44	
Predictor						
BLAT	-0.40	(6)	-		-	
BIM12	-0.25	(1)*	-0.53	(1)*	-0.77	(1)*
DBIM10	0.40	(2)*	-		-	
C4U 1	0.39	(3)*	1.45	(2)*	1.83	(2)*
C4U 2	-		0.72	(3)*	0.82	(5)
C4U20	1.27	(5)	-		-	
C4V12	-		1.03	(7)	1.34	(6)
C4V15	1.04	(4)	-		-	
C7U 7	-		-1.32	(4)*	-1.98	(3)*
S27U24	-		1.88	(6)	-	
S27V 1	-		-0.78	(5)	-0.98	(4)
Cases	588		457		346	
R**2	0.42		0.52		0.59	

The verification of the SCRN predictor equations derived from the smaller dependent-case subset (applied to both the dependent-case and independent-case subsets) is summarized relative to homogeneous samples of persistence and JTWC forecasts in Table 11. For ease of comparison, the verification results of the SCRN

TABLE 11

Verification of persistence, JTWC and regression-derived (SCRN) 24, 48 and 72 h forecasts of western North Pacific tropical storm intensity (kt) using land-filter and homogeneous data from the complete dependent data set (CDS), the dependent-case subset (DCS), and the independent-case subset.

Complete Dependent Data Set - CDS SCRN Predictor Eqns

		Persistence		JTWC		Regression	
	Cases	Av Abs Error	Std Dev	Av Abs Error	Std Dev	Av Abs Error	Std Dev
24 h	886	17.2	13.7	13.1	11.3	13.0	11.2
48 h	651	28.1	18.8	21.3	16.6	19.5	14.8
72 h	462	33.7	23.9	24.5	19.0	21.3	16.1

Dependent-Case Subset - DCS SCRN Predictor Eqns

		Persistence		JTWC		Regression	
	Cases	Av Abs Error	Std Dev	Av Abs Error	Std Dev	Av Abs Error	Std Dev
24 h	587	17.0	13.3	13.0	11.3	12.4	10.8
48 h	439	28.1	19.3	21.3	16.9	18.9	14.6
72 h	312	33.0	24.5	24.3	19.1	20.6	16.5

Independent-Case Subset - DCS SCRN Predictor Eqns

		Persistence		JTWC		Regression	
	Cases	Av Abs Error	Std Dev	Av Abs Error	Std Dev	Av Abs Error	Std Dev
24 h	299	17.7	14.3	13.2	11.4	14.5	11.8
48 h	212	28.2	21.2	21.2	15.9	20.2	14.9
72 h	150	35.1	22.5	24.8	18.7	22.1	16.8

predictor equations derived from the complete dependent set are also repeated from Table 9. The similar characteristics (mean absolute errors and standard deviations)

between the homogeneous samples of persistence and JTWC forecasts associated with the complete dependent set and the dependent-case subset implies the dependent-case subset is a representative sample of the complete dependent set. As expected, regression equations derived from the smaller dependent-case subset perform better than the equations derived from the complete dependent sample (average absolute errors of 12.4, 18.9 and 20.6 kt versus 13.0, 19.5 and 21.3 kt for 24, 48 and 72 h, respectively). This fictitious improvement is attributed to either a dependent-case sample size that is too small for proper development of the regression equations, or F-to-enter and R^2 criteria that are too lenient for properly restricting the predictor selection. However, when the dependent-case equations are applied to the independent-case subset, the good performance suggested by the dependent-case results is not sustained. Nevertheless the performance is better (smaller average absolute errors) and is more consistent (smaller standard deviations) than JTWC official forecasts at 48 and 72 h. For example, the mean absolute errors in this independent sample are 20.2 and 22.1 kt versus 21.2 and 24.8 kt for JTWC.

3. **Subsets stratified by previous 12 h intensity**

Recall that the 12 h old best track intensity is the first predictor chosen in the SCRN predictor intensity-change equations for the complete dependent set at all three forecast periods (Table 4). Therefore, the 1216 data cases are subdivided into terciles based upon the 12 h old best track intensity. This is a common practice in that conventional-predictor forecast schemes currently in use at the operational forecast centers are stratified by intensity. The frequency distribution of 12 h old best track intensity values and the tercile cut-points are illustrated in Fig. 4.2. The stratification scheme used to subdivide the data cases into weak, moderate and strong subsets based on 12 h old intensity is illustrated in Table 12. An exact division into three equal-size categories is not possible because intensity values are recorded to the nearest 5 kt.

The basic approach in Fig. 4.1 (Section A above) is applied to each of the three tercile subsets. The results of the screening process on each category of potential predictors (CONV, WIND and SHEAR) selected for each tercile subset and each forecast interval are indicated in Tables 13, 14 and 15. The values in the nine columns on each table indicate the order in which the screened predictors were selected as the next dominant predictor for the associated data subset and forecast interval. As with the smaller dependent-case sample (Section B.2 above), more predictors generally are selected for each equation when the data are stratified into terciles (smaller sample

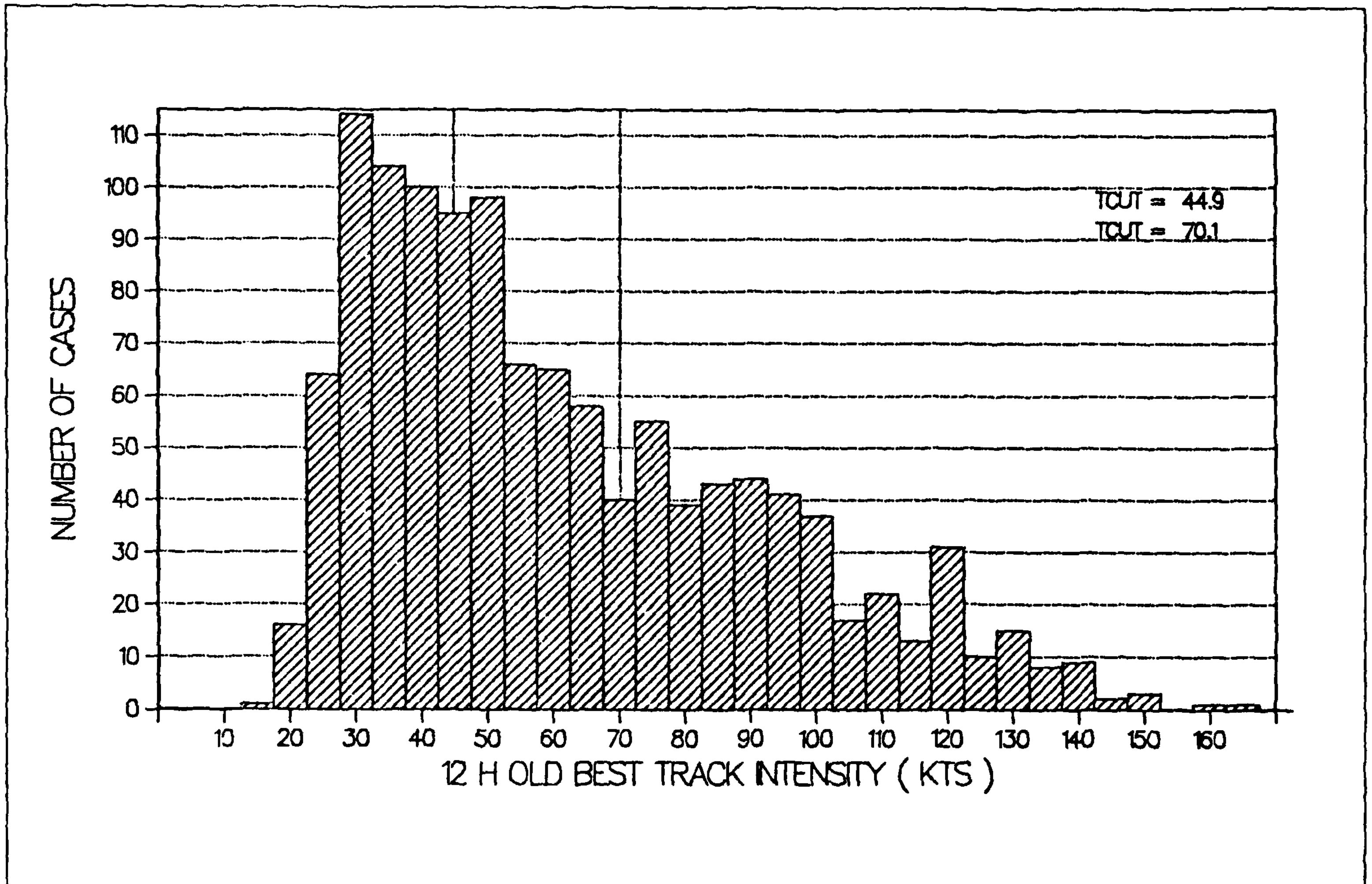


Figure 4.2 Histogram of 12 h old best track intensity (kt) with tercile cut-points between weak-moderate and moderate-strong storms.

TABLE 12

Stratification scheme for the tercile subsets with data stratified according to previous 12 h best track intensity (kt).

Class	Cases	Intensity
Weak	399	I < 45 kt
Moderate	422	45 kt < I < 70 kt
Strong	391	70 kt < I

sizes) than are selected using the complete dependent set. This may be misleading, as it was with the smaller dependent-case subset, when independent cases are examined.

TABLE 13

Conventional predictors (CONV) selected after screening regression on 24, 48 and 72 h best track intensity change (kt). Stratified data sets are based on 12 h old best track intensity (kt). The numbers indicate the order in which the predictors were selected.

	WEAK			MODERATE			STRONG		
PREDICTORS	24h	48h	72h	24h	48h	72h	24h	48h	72h
DAYJUL		2	1		4	2			
BLAT					5	4	3	3	3
DBIM10	1	1		1	1		1	2	
BLTM12			2	2	3	3	4	5	
BLNM12	2			3		1	5	4	2
BIOO					2		2	1	1

The screening on CONV predictors illustrates several points (Table 13). Selection of past 12 h intensity change as either the first or second 24 and 48 h conventional predictor for all three tercile subsets implies extrapolation of the intensity trend is useful as a technique for the shorter range intensity forecast, but not the 72 h forecast. The current intensity is the first (48 and 72 h) or second (24 h) predictor selected for all forecast periods using the strong tercile. This suggests persistence is a useful parameter in the forecast of stronger storms.

The results of screening with WIND predictors using the tercile subsets (Table 14) and the complete dependent set (Table 5) may be compared. A total of 57 WIND predictors are selected for the nine equations using tercile subsets, as opposed to only 15 predictors selected for the three complete dependent set equations. Six of the WIND predictors (C2U14, C2V25, C4U 2, C4U16, C7V13 and C7V14) selected using the complete dependent set are not selected using the intensity-stratified subsets. This observation is surprising because C4U 2 was the second predictor selected in the 24 and 72 h equations (fourth in the 48 h equation) for the WIND predictor screening using the complete dependent set (Table 5). Furthermore, C4U 2 entered all three SCRN predictor equations using the complete dependent set (Table 8).

The results of the screening with SHEAR predictors using the tercile subsets (Table 15) may be compared with the complete dependent set (Table 6). A total of 58 SHEAR predictors are selected in the nine equations for tercile subsets as opposed to 17 SHEAR predictors selected in the three equations for the complete dependent set. Five of the SHEAR predictors (S47U 4, S24U 2, S27V 4, S27V19 and S27V22) selected using the complete dependent set are not selected using the intensity-stratified subsets.

The SCRN predictor equations for the 24, 48 and 72 h intensity are illustrated in Tables 16, 17 and 18, which correspond to the weak, moderate and strong subsets, respectively. Analysis of the equations selected for the weak tercile (Table 16) indicates that lower layer wind-shear EOF coefficients are selected first (S47U11) in the 72 h forecast equation and second (S47U 1) in the 24 h forecast equation. Selection of past 12 h change in intensity as a predictor in the 24 and 48 h (but not 72 h) equation suggests the usefulness of extrapolation in the short-term forecast with weak tercile storms.

TABLE 14

Wind EOF coefficient predictors (WIND) selected after screening regression on 24, 48 and 72 h best track intensity change (kt). Stratified data sets are based on 12 h old best track intensity (kt). The numbers indicate the order in which the predictors were selected.

	WEAK			MODERATE			STRONG		
	24h	48h	72h	24h	48h	72h	24h	48h	72h
PREDICTORS									
C2U 1							1	1	
C2U 6									8
C2U 8		7							
C2U 9							10		
C2U13			8						
C2U17									2
C2U19		4		7		5			
C2U20	2								
C2V 1		5		4	5				1
C2V18								7	
C2V23							4	3	
C2V24							9		
C4U 1	1	1		1	1	2			
C4U 5	8								
C4U10								9	
C4U14	6								
C4U20				3	4				
C4U22			4						
C4U25				6					
C4V 2									6
C4V 7	3	2							
C4V 8					10				
C4V11						9			
C4V12			1						
C4V15	4							2	
C4V16	10								5
C4V19		3	3	8					
C4V20								8	
C4V23							6		
C4V24							3		

TABLE 14

(cont'd.)

C7U 2				2					
C7U 3	7		5						
C7U 4									3
C7U 5					6	8		4	
C7U 6						7			
C7U 7			6	5	2	1			
C7U 9		8							
C7U10								5	
C7U11			2						
C7U12									9
C7U13									10
C7U14	5			10	7	4			
C7U17					8				7
C7U20						5			
C7U24			9						
C7V 1						6			
C7V 3					3	3			
C7V 5									4
C7V 8			7						
C7V 9		6			8		2		
C7V11				9					
C7V15	9	9							
C7V16							7	6	
C7V18								10	
C7V20		10							
C7V21						10			
C7V22							8		

TABLE 15

Vertical wind-shear EOF coefficient predictors (SHEAR) selected after screening regression on 24, 48 and 72 h best track intensity change (kt). Stratified data sets are based on 12 h old best track intensity (kt). The numbers indicate the order in which the predictors were selected.

	WEAK			MODERATE			STRONG		
PREDICTORS	24h	48h	72h	24h	48h	72h	24h	48h	72h
S47U 1	1	2		1	3			3	
S47U 2							2	2	
S47U 3		7						8	
S47U 8		9							
S47U10						9			
S47U11	6	1	1					6	
S47U12			8						
S47U14			2						
S47U15					10				
S47U16							8	4	2
S47U17				8	1	10			
S47U22						2			
S47U24				2	9				
S47V 3				3	4				
S47V 6			4						
S47V 7	10								
S47V12							5		
S47V15									9
S47V16					5	5			
S47V18	8								
S47V19									7
S47V21			7						
S47V25						8			
S24U 1				7			1	1	
S24U 3	2								
S24U 4	5								
S24U17		8							
S24U20								10	10
S24U21			3						
S24U22						1			

TABLE 15

(cont'd.)

S24U23				5					
S24U24				9					
S24V 2								7	
S24V 3					2	7			
S24V 4									4
S24V 8	7						9	5	
S24V 9					6				
S24V14						3			
S24V15	3								
S24V18			6						
S24V20							6		
S24v24	9	3			8	6			
S24V25	4					4			
S27U 1		5							
S27U 4									5
S27U 7							3	3	
S27U 9									6
S27U10									8
S27U18				6				9	
S27U23			5						
S27U24		6							
S27V 1				4					
S27V 9					7				
S27V11							4		
S27V16		4							
S27V17									1
S27V20				10		7			
S27V24		10							

TABLE 16

Regression equations for the change in intensity (kt) at 24, 48 and 72 h using data stratified by 12 h old best track intensity (WEAK tercile). Values in parentheses indicate the order in which SCRN predictors were selected. The coefficients of multiple determination (R**2) are shown.

	Forecast Interval		
	24 h	48 h	72 h
Y-Intercept	9.00	22.22	15.58
Predictor			
DAYJUL	-	-	0.07 (10)
DBIM10	0.54 (1)	1.06 (1)	-
C2U 8	-	-1.38 (10)	-
C2U19	-	2.66 (2)	-
C2U20	2.41 (3)	-	-
C2V 1	-	-1.34 (6)	-
C4U 1	-	-1.36 (2)	-
C4U14	-0.94 (5)	-	-
C4U22			4.92 (5)
C4V 7	0.61 (6)	2.33 (4)	-
C4V12	-	-	2.79 (2)
C4V15	0.78 (7)	-	-
C4V19	-	2.34 (3)	2.76 (4)
C7U 3	-	-	0.83 (9)
C7U 7	-	-	-2.10 (7)
C7U 9	-	-2.14 (7)	-
C7U11	-	-	-1.74 (3)
C7U14	-0.47	-	-
C7V 8	-	-	1.66 (6)
C7V 9	-	-1.45 (9)	-
C7V15	-	1.39 (8)	-
S47U 1	0.46 (2)	-	-
S47U11	-	-	1.83 (1)
S47U14	-	-	-2.29 (8)
S24V 8	-0.41 (9)	-	-
S24V15	-0.96 (4)	-	-
S24V24	-0.68 (10)	-	-
Cases	249	202	167
R**2	0.42	0.48	0.43

TABLE 17

Regression equations for the change in intensity (kt) at 24, 48 and 72 h using data stratified by 12 h old best track intensity (MODERATE tercile). Values in parentheses indicate the order in which SCRN predictors were selected. The coefficients of multiple determination (R**2) are shown.

	Forecast Interval					
	24 h		48 h		72 h	
Y-Intercept	4.12		6.56		-45.63	
Predictor						
BLTM12	-		-		0.37	(10)
DBIM10	0.58	(1)	-		-	
C2U19	1.33	(9)	-		-	
C2V 1	-0.64	(4)	-1.54	(5)	-	
C4U 1	0.83	(2)	1.84	(1)	1.16	(2)
C4U20	2.02	(5)	3.25	(4)	-	
C4U25	1.79	(7)	-		-	
C4V 8	-		1.23	(8)	-	
C7U 2	0.39	(3)	-		-	
C7U 5	-		0.81	(6)	-	
C7U 6	-		-		-1.35	(7)
C7U 7	-0.73	(6)	-2.17	(2)	-2.74	(1)
C7U14	-		-1.43	(10)	-3.15	(4)
C7U17	-		2.19	(9)	-	
C7U20	-		-		-3.45	(5)
C7V 1	-		-		1.24	(6)
C7V 3	-		-1.23	(3)	-0.62	(3)
S47U17	0.94	(10)	2.13	(7)	-	
S24U22	-		-		-2.61	(9)
S24U23	-1.26	(8)	-		-	
S24V25	-		-		2.66	(8)
Cases	320		249		185	
R**2	0.43		0.47		0.49	

TABLE 18

Regression equations for the change in intensity (kt) at 24, 48 and 72 h using data stratified by 12 h old best track intensity (STRONG tercile). Values in parentheses indicate the order in which SCRN predictors were selected. The coefficients of multiple determination (R**2) are shown.

	Forecast Interval					
	24 h		48 h		72 h	
Y-Intercept	-29.99		-7.27		-84.56	
Predictor						
BLAT	-		-1.12	(7)	-1.16	(4)
BLNM12	0.38	(6)	0.43	(10)	1.06	(2)
BI00	-0.29	(3)	-0.50	(1)	-0.63	(1)
DBIM10	0.35	(2)	0.27	(6)	-	
C2U 1	0.61	(1)	0.75	(2)	-	
C2U 6	-		-		1.87	(7)
C2V 1	-		-		1.83	(3)
C4U10	-		1.77	(8)	-	
C4V 2	-		-		-1.50	(8)
C4V15	-		1.63	(4)	-	
C4V20	-		-2.46	(5)	-	
C4V23	0.98	(10)	-		-	
C4V24	1.34	(5)	-		-	
C7U17	-		-		3.94	(6)
C7V 5	-		-		-1.50	(5)
C7V 9	0.75	(4)	-		-	
C7V16	0.90	(9)	-		-	
C7V18	-		-1.76	(3)	-	
S47U 2	-0.35	(8)	-		-	
S24V 4	-		-		1.30	(9)
S27U 4	-		-		1.16	(10)
S24U 7	0.62	(7)	1.28	(8)	-	
Cases	318		233		160	
R**2	0.36		0.51		0.61	

SHEAR predictors are not among the first six predictors chosen for any of the moderate or strong tercile equations (Tables 17 and 18). This might be physically relevant in that a weak storm will not develop with large environmental shear, but if the storm develops to more than 45 kt (moderate or strong tercile case), vertical wind shear is not a significant factor in deciding further intensity changes.

Notice that the explained variance values are largest at 48 h (0.51), and especially at 72 h (0.61), in the strong tercile equations. By contrast, the explained variance is lowest (0.36) at 24 h in the strong category.

The verifications of the equations for the three terciled subsets are illustrated in Table 19. The verification of the complete dependent data set is repeated from Table 9 for comparison. The average absolute error for the 24 h forecast is smallest (9.3 kt) for the weak tercile and largest (13.2 kt) with the strong tercile equations. The average absolute errors for the 48 h equations are comparable for all three terciles (16.4, 17.6 and 15.4 kt). The average absolute error for the 72 h equation is smallest for the strong tercile equation (14.7 kt) and largest for the weak tercile (18.7 kt).

The weighted-average absolute error of the regression-derived intensity is computed for each forecast period as

$$AAE_{mean} = \frac{(N_w AAE_w + N_m AAE_m + N_s AAE_s)}{N_{total}} , \qquad (4.2)$$

where AAE indicates the average absolute error and N is the sample size. The subscripts refer to the particular data set; i.e., the complete dependent set (total) or a subset of the complete set statified by 12 old best track intensity ('w' indicates the weak, 'm' indicates the moderate, and 's' indicates the strong tercile).

TABLE 19

Verification of 24, 48 and 72 h
tropical storm intensity forecasts (kt)
for WEAK, MODERATE and STRONG terciles
(stratified by 12 h old intensity)
using SCRN predictor equations.

Complete dependent set - SCRN predictors

	Cases	Avg Abs Error	Std Dev
24 h	886	13.0	11.2
48 h	651	19.5	14.8
72 h	462	21.3	16.1

WEAK tercile subset - SCRN predictors

	Cases	Avg Abs Error	Std Dev
24 h	249	9.3	7.2
48 h	190	16.4	11.7
72 h	148	18.7	15.5

MODERATE tercile subset - SCRN predictors

	Cases	Avg Abs Error	Std Dev
24 h	319	11.7	10.7
48 h	233	17.6	13.6
72 h	163	17.9	13.7

STRONG tercile subset - SCRN predictors

	Cases	Avg Abs Error	Std Dev
24 h	318	13.2	10.5
48 h	228	15.4	12.6
72 h	151	14.7	11.8

The verifications of the intensity-stratified equations are outlined in Table 20. Verifications of the homogeneous persistence, JTWC and complete dependent set forecasts are repeated from Table 9 for comparison. The weighted-average absolute errors indicate that the intensity-stratified equations based on dependent cases perform better than all other forecast schemes over all forecast intervals. Student-T tests confirm that the tercile subsets are significantly better (95% confidence level) than the JTWC official forecast at 48 and 72 h.

TABLE 20

Verification of 24, 48 and 72 h tropical storm intensity forecasts (kt): (1) persistence, (2) JTWC, (3) regression (complete set) and (4) regression (stratified subsets, based on 12 old best track intensity).

Persistence Forecast

	Cases	Avg Abs Error	Std Dev
24 h	886	17.2	13.7
48 h	651	28.1	19.8
72 h	462	33.7	23.9

JTWC Forecast

	Cases	Avg Abs Error	Std Dev
24 h	886	13.1	11.3
48 h	651	21.3	16.6
72 h	462	24.5	19.0

Regression-Derived Forecast
Screened predictors, Unstatified Data

	Cases	Avg Abs Error	Std Dev
24 h	886	13.0	11.2
48 h	651	19.5	14.8
72 h	462	21.3	16.1

Regression-Derived Forecast
Screened Predictors, Stratified Data

	Cases	WT Avg Abs Error	Std Dev
24 h	886	11.6	---
48 h	651	16.5	---
72 h	462	17.1	---

V. SUMMARY AND RECOMMENDATIONS

This study is the first step in the development of an enhanced objective technique for predicting 24, 48 and 72 h intensity of tropical cyclones in the western North Pacific region. The eventual goal is to develop an effective aid for the Joint Typhoon Warning Center (JTWC) to forecast tropical storm intensity, particularly at 48 and 72 h.

The EOF coefficients of zonal and meridional components of the environmental wind at 250, 400 and 700 mb (Wilson, 1984) and wind shear from 250 to 400, from 400 to 700, and from 250 to 700 mb (Meanor, 1987) are considered as potential predictors. Additional predictors include conventional storm-related parameters, such as date, intensity, motion and position. The 1216 cases in this study are 12 h data for western North Pacific tropical cyclones from 1979 to 1983. The basic methodology involves the following four steps:

- Select a data set, i.e., complete dependent set; independent-case/dependent-case subsets; or subsets stratified by 12 h old intensity;
- Screen predictors using stepwise regression analysis to select the dominant predictors;
- Generate regression equations using stepwise regression analysis and the screened predictors to generate regression equations; and
- Verify the equations relative to the performance of the Joint Typhoon Warning Center official forecast.

When the basic methodology is applied to a complete set (1216 cases), the regression equations using only conventional predictors are slightly improved by inclusion of synoptic forcing fields represented by the EOF coefficients. Furthermore, the regression equations perform slightly better than the JTWC official forecasts.

When the equations generated using a smaller dependent-case subset (811 cases) are applied to the dependent-case subset, similar results are observed. Relative to a homogeneous sample of JTWC official forecasts, the regression equations developed using the dependent-case subset show progressively improved performance with increasing forecast interval. Despite a slight increase in the average absolute error at all forecast intervals when the dependent-case equations are applied to the independent-case subset, the performance of these equations is still comparable to a homogeneous sample of JTWC official forecasts.

When the basic method is applied to subsets stratified by 12 h old intensity, the regression equations perform better than the JTWC official forecast at all forecast intervals. These equations are significantly better (95% confidence) than the JTWC official forecast at 48 and 72 h.

These results suggest that the official JTWC tropical storm intensity forecasts can be enhanced by application of statistical regression analysis techniques. The performance of the existing techniques based on conventional storm-related predictors can be progressively improved by using:

- regression equations based on selected screened predictors drawn from EOF coefficient predictors of wind and vertical wind shear; and
- regression equations developed from tercile subsets (the cases are statified by 12 h old intensity) and selected screened predictors.

The preliminary success of the screened regression equations, particularly those developed using the stratified case subsets, suggests EOF coefficients of the wind and the vertical wind-shear fields should be computed routinely using current data. With the EOF coefficient predictors routinely available, these objective techniques could be tested in an operational environment using independent cases. To further enhance these objective techniques, the predictive ability of other synoptic or remotely sensed parameters should be investigated.

LIST OF REFERENCES

Dixon, W. J., and M. B. Brown, 1985: BMDP Statistical Software Manual, University of California Press, Los Angeles, CA, 734 pp.

Dvorak, V. F., 1975: Tropical cyclone intensity analysis and forecasting from satellite imagery. *Mon. Wea. Rev.*, **103**, 420-430.

____, 1982: Tropical cyclone intensity analysis and forecasting using enhanced infrared imagery. Applications Laboratory Training Notes, National Oceanic and Atmospheric Administration, National Earth Satellite Service, U. S. Department of Commerce, Washington, D. C., 22 pp.

Elsberry, R. L., G. G. Coltrane and P. L. Krueger, 1975: Statistical forecasts of 24, 48 and 72 h typhoon and tropical storm intensity changes. *J. Appl. Meteor.*, **14**, 445-451.

George, J. E., and W. M. Gray, 1976: Tropical cyclone motion and surrounding storm relationships. *J. Appl. Meteor.*, **15**, 1252-1264.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity for the North Atlantic Basin. *NOAA Technical Memorandum*, NWS NHC 10, National Hurricane Center, Miami, FL, 22 pp.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather predictions. Scientific Report 1, Statistical Forecasting Project, Department of Meteorology, Massachusetts Institute of Technology, Cambridge, MA, 48 pp.

Meanor, D. H., 1987: Vertical wind shear as a predictor of tropical cyclone motion. M. S. Thesis, Naval Postgraduate School, Monterey, CA, 74 pp.

Merrill, R. T, 1987: An experiment in the statistical prediction of tropical cyclone intensity change. Preprints, 17th Annual Conf. on Hurr. and Trop. Meteor., Miami, FL, April 7-10, 1987, American Meteor. Soc., Boston, MA, 302-304.

Peak, J. E., and R. L. Elsberry, 1987: Selection of optimal tropical cyclone motion guidance using an objective classification tree methodology. To appear in *Mon. Wea. Rev.*

Schott, T. B., 1985: Applications of wind empirical orthogonal functions in tropical cyclone motion studies. M. S. Thesis, Naval Postgraduate School, Monterey, CA, 99 pp.

Shaffer, A. R., 1982: Typhoon motion forecasting using empirical orthogonal function analysis of the synoptic forcing. M. S. Thesis, Naval Postgraduate School, Monterey, CA, 150 pp.

Shaffer, A. R., and R. L. Elsberry, 1982: A statistical climatological tropical cyclone track prediction technique using an EOF representation of the synoptic forcing. *Mon. Wea. Rev.*, **110**, 1945-1954.

U. S. Naval Command Center/Joint Typhoon Warning Center, 1985: *Annual Tropical Cyclone Report*, COMNAVMARIANAS Box 17, FPO San Francisco, CA, 274 pp.

Wilson, W. E., 1984: Forecasting of tropical cyclone motion using an EOF representation of wind forcing, M. S. Thesis, Naval Postgraduate School, Monterey, CA, 86 pp.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Chairman (Code 63Rd) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-5000	1
4.	Professor R. L. Elsberry (Code 63Es) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-5000	5
5.	Professor F. R. Williams (Code 63Wf) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Capt Edward L. Weniger 28WS/DN APO NY 09127	1
7.	Program Manager (AFIT/CIR) Air Force Institute of Technology Wright-Patterson AFB, OH 45433	1
8.	Commander Air Weather Service Scott AFB, IL 62225	1
9.	Air Weather Service Technical Library Scott AFB, IL 62225	2
10.	Commander Air Force Global Weather Central Offutt AFB, NE 68113	1
11.	Director Naval Oceanography Division Naval Observatory 34th and Massachusetts Avenue NW Washington, DC 20390	1

12.	Commander Naval Oceanography Command NSTL Station Bay St. Louis, MS 39522	1
13.	Chief of Naval Research 800 N. Quincy Street Arlington, VA 22217	1
14.	Commanding Officer Naval Environmental Prediction Research Facility Monterey, CA 93943-5000	1
15.	Dr. Ted Tsui Naval Environmental Prediction Research Facility Monterey, CA 93943-5000	1
16.	Commanding Officer Fleet Numerical Oceanography Center Monterey, CA 93943-5000	1
17.	Mr. Charles Mauch Fleet Numerical Oceanography Center Monterey, CA 93943-5000	1
18.	Commanding Officer Naval Oceanography Command Center, Guam Box 17, COMNAVMARIANAS FPO San Francisco, CA 96630	1
19.	Director Joint Typhoon Warning Center Box 17, COMNAVMARIANAS FPO San Francisco, CA 96630	1
20.	Commanding Officer Naval Western Oceanography Center Box 113 Pearl Harbor, HI 96818	1
21.	Commanding Officer Naval Eastern Oceanography Center Naval Air Station Norfolk, VA 23511	1
22.	Chairman Oceanography Department U.S. Naval Academy Annapolis, MD 21402	1
24.	Mr. and Mrs. Edward W. Weniger 368 Hazel Avenue Feasterville, PA 19047	1

END

10-87

DTIC